Available online at www.sciencedirect.com
ScienceDirect
Remote Sensing of Environment 112 (2008) 1658–1677
Remote Sensing of Environment
www.elsevier.com/locate/rse

Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information

J.A. Blackard [a], M.V. Finco [b], E.H. Helmer [c], G.R. Holden [d], M.L. Hoppus [e], D.M. Jacobs [f], A.J. Lister [e], G.G. Moisen [a,*], M.D. Nelson [d], R. Riemann [e], B. Ruefenacht [b], D. Salajanu [f], D.L Weyermann [g], K.C. Winterberger [h], T.J. Brandeis [f], R.L. Czaplewski [i], R.E. McRoberts [d], P.L. Patterson [a], R.P. Tymcio [a]

[a] *Rocky Mtn. Research Station, 507 25th Street, Ogden, UT 84401, United States*
[b] *Remote Sensing Applications Center, 2200 W 2300 S, Salt Lake City, UT 84119, United States*
[c] *International Institute of Tropical Forestry, Jardín Botánico Sur, 1201 Calle Ceiba, Río Piedras, 00926, Puerto Rico*
[d] *North Central Research Station, 1992 Folwell Ave, St. Paul, MN 55108, United States*
[e] *Northeastern Research Station, 11 Campus Blvd, Newtown Square, PA 19073, United States*
[f] *Southern Research Station, 4700 Old Kingston Pike, Knoxville, TN 37919, United States*
[g] *Pacific Northwest Research Station, 1221 SW Yamhill, Portland, OR 97205, United States*
[h] *Pacific Northwest Research Station, 3301 C St, Anchorage, AK 99503, United States*
[i] *Rocky Mtn Research Station, 240 W Prospect Rd, Fort Collins, CO 80526, United States*

Received 20 December 2005; received in revised form 24 August 2007; accepted 24 August 2007

Abstract

A spatially explicit dataset of aboveground live forest biomass was made from ground measured inventory plots for the conterminous U.S., Alaska and Puerto Rico. The plot data are from the USDA Forest Service Forest Inventory and Analysis (FIA) program. To scale these plot data to maps, we developed models relating field-measured response variables to plot attributes serving as the predictor variables. The plot attributes came from intersecting plot coordinates with geospatial datasets. Consequently, these models serve as mapping models. The geospatial predictor variables included Moderate Resolution Imaging Spectrometer (MODIS)-derived image composites and percent tree cover; land cover proportions and other data from the National Land Cover Dataset (NLCD); topographic variables; monthly and annual climate parameters; and other ancillary variables. We segmented the mapping models for the U.S. into 65 ecologically similar mapping zones, plus Alaska and Puerto Rico. First, we developed a forest mask by modeling the forest vs. nonforest assignment of field plots as functions of the predictor layers using classification trees in See5©. Secondly, forest biomass models were built within the predicted forest areas using tree-based algorithms in Cubist©. To validate the models, we compared field-measured with model-predicted forest/nonforest classification and biomass from an independent test set, randomly selected from available plot data for each mapping zone. The estimated proportion of correctly classified pixels for the forest mask ranged from 0.79 in Puerto Rico to 0.94 in Alaska. For biomass, model correlation coefficients ranged from a high of 0.73 in the Pacific Northwest, to a low of 0.31 in the Southern region. There was a tendency in all regions for these models to over-predict areas of small biomass and under-predict areas of large biomass, not capturing the full range in variability. Map-based estimates of forest area and forest biomass compared well with traditional plot-based estimates for individual states and for four scales of spatial aggregation. Variable importance analyses revealed that MODIS-derived information could contribute more predictive power than other classes of information when used in isolation. However, the true contribution of each variable is confounded by high correlations. Consequently, excluding any one class of variables resulted in only small effects on overall map accuracy. An estimate of total C pools in live forest biomass of U.S. forests, derived from the nationwide biomass map, also compared well with previously published estimates.

Keywords: Forest biomass; MODIS; Classification and regression trees; Forest probability; Carbon; FIA

* Corresponding author. USDA Forest Service, Rocky Mountain Research Station, 507 25th Street, Ogden, UT 84401, United States, Tel.: +1 801 625 5384; fax: +1 801 625 5723.
E-mail address: gmoisen@fs.fed.us (G.G. Moisen).

0034-4257/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2007.08.021

1. Introduction

The Forest Inventory and Analysis (FIA) program of the USDA Forest Service collects data annually on the status and trends in forested ecosystems nationwide. These inventory data support estimates of forest population totals over large geographic areas, (Scott et al., 2005). Regional maps of forest characteristics would make these extensive forest resource data more accessible and useful to a larger and more diverse audience. Important applications of such maps include broad-scale mapping and assessment of wildlife habitat; documenting forest resources affected by fire, fragmentation, and urbanization; identifying land suitable for timber production; and locating areas at high risk for plant invasions, or insect or disease outbreaks. Thus, there is a need to produce and distribute geospatial data of forest attributes, complementing FIA inventory data.

Total aboveground live biomass is a forest characteristic of particular interest. Forest soils and woody biomass hold most of the carbon in Earth's terrestrial biomes (Houghton, 1999). Land-use change, mainly forest burning, harvest, or clearing for agriculture, may compose 15 to 40% of annual human-caused emissions of carbon to the atmosphere, and terrestrial ecosystems, mainly through forest growth and expansion, absorb nearly as much carbon annually. However, estimates of land-atmosphere carbon fluxes, and the net of expected future ones, have the largest uncertainties in the global atmospheric carbon budget, which adds to uncertainties about future levels and impacts of greenhouse gasses (GHGs) in the atmosphere (Houghton, 2003; Prentice et al., 2001).

Consequently, the levels, mechanisms and spatial distribution of forest land-atmosphere C fluxes are an important focus for reducing uncertainties in the global C budget (Fan et al., 1998; Holland et al., 1999; Pacala et al., 2001; Schimel et al., 2001). Ecosystem process models that are physiologically-based, and that use satellite image-derived indices of photosynthesis, have permitted unprecedented global assessments of ecosystem productivity and carbon sinks at a spatial resolution of 0.5° (Nemani et al., 2003; Potter et al., 2003). The mechanistic nature of these models identifies how observed patterns in ecosystem productivity may relate to climate and atmospheric changes (Nemani et al., 2003). However, validating atmospheric and ecosystem model estimates of net forest C fluxes, and quantifying the C fluxes associated with changes in land use, which dominate these fluxes over longer time periods, requires spatially extensive data on forest C pools and net fluxes. Maps of forest biomass permit spatially explicit estimates of forest carbon storage and net fluxes from land-use change.

Our objectives here are to 1) produce a spatially explicit dataset of aboveground live forest biomass from ground measured inventory plots, at a 250-m cell size, for the conterminous U.S., Alaska and Puerto Rico; 2) evaluate model performance and spatially depict uncertainty in the dataset; 3) explore the relative contribution of the many predictor layers to the biomass models; and 4) use the resulting dataset to estimate aboveground live forest biomass and implied carbon storage for this area. We also describe a national geospatial predictor database that supported the mapping and how we standardized national FIA data, developed predictive models, and assessed model error.

2. Methods

2.1. Data

2.1.1. Response variables

The US Forest Service FIA program inventories the Nation's forests via a network of ground-based inventory plots in which forest structure and tree species composition are measured to produce estimates of forest attributes like basal area by species, total volume, and total biomass. Plots are located with an intensity of about one plot per 2400 ha. Although the program historically collected data periodically (every 5 to 20 years) for each state in the country, it recently shifted to an annual rotating panel system. This new system samples 10 to 20% of each state's plot network annually (Bechtold & Patterson, 2005). This study used a mixture of annual and historic periodic data to ensure enough training plots in all parts of the country, with dates of collection ranging between 1990 and 2003. The advantages of modeling response variables collected from a probabilistic sample (such as FIA's plot network) over those collected from a purposive sample are explored in Edwards et al. (2006).

The FIA program observes, measures, and predicts many forest attributes on each plot (Miles et al., 2001). This nationwide biomass mapping effort modeled two of these plot-level response variables: a binary forest/nonforest classification and above-ground live forest biomass. According to FIA definitions, *forest land* is at least 0.405 ha in size, has a minimum continuous canopy width of 36.58 m with at least 10% stocking, and has an understory undisturbed by a nonforest land use like residences or agriculture. *Aboveground live biomass* includes biomass in live tree bole wood, stumps, branches and twigs for trees 2.54-cm diameter or larger and is derived from region- or species-specific allometric equations.

2.1.2. National geospatial predictor layers

A nationwide geospatial dataset of layers of predictor variables, also called the national geospatial predictor layer database, was assembled for use in the biomass models. The data layers included satellite imagery and predicted land-cover from Moderate Resolution Imaging Spectro-radiometer (MODIS) (Justice et al., 2002), Landsat Thematic Mapper image-derived National Land Cover Dataset (NLCD92, Vogelmann et al., 2001), raster climate data, and topographic variables. Datasets with native spatial resolutions other than 250 m were resampled with a nearest neighbor procedure if categorical, and a bilinear interpolation procedure if continuous. The 250-m spatial resolution of the predictor dataset has two origins. First, the coarser spatial scale of MODIS would be practical given the national extent of the project, and the MODIS sensor bands 1 and 2 are available at that spatial resolution. As a result, MODIS vegetation index data are available with 250-m pixel sizes. Secondly, we expected that coarser image data would have scaling advantages when working with passive optical imagery, as we discuss later.

Data from MODIS for the year 2001 included all land surface reflectance bands (Vermote & Vermueulen, 1999) (MOD 09v003) from three 8-day image composites at 500-m resolution (beginning Julian days 097, 225, 321), three 16-day vegetation index (VI) composites (Huete et al., 2002) (MOD 13v003) at 250-m resolution over the same three compositing periods, and percent tree cover (MOD 44) at 500-m resolution for 2001 (Hansen et al., 2003). The compositing periods represented early, peak, and leaf-off phenological conditions in the continental United States. For Puerto Rico, persistent cloudiness necessitated data from dry-season MODIS image compositing periods, including six periods from 2001–2003. The MOD 09 8-day image composites use a minimum-blue criterion to select for clearest conditions (Vermote & Vermueulen, 1999). The compositing algorithm for MOD 13 VI data first selects clear pixels over the compositing period with the MODIS cloud mask. A pixel-level fit to a bidirectional reflectance distribution function (BRDF) then estimates a near-nadir reflectance for each band for calculating VI values. If fewer than five pixels are clear over the compositing period, then the algorithm selects a clear pixel based on viewing angle. Otherwise, the algorithm selects the pixel with the maximum Normalized Difference Vegetation Index (NDVI) (Huete et al., 2002). We performed no additional image compositing or cloud filling for continental U.S. imagery. Some cloudy areas were masked from the Puerto Rico composites and filled with appropriate composite imagery from other dates.

Landsat image-based land cover for the conterminous U.S. (Vogelmann et al., 2001) and Puerto Rico (Helmer et al., 2002) provided data on proportional cover of forest, shrubland, wetland and urban/barren lands (Puerto Rico only). These 30-m components of the national geospatial predictor data used focal functions to summarize the land cover class proportions within a 9×9 moving window and subsequently resampled the data to 250-m with bilinear resampling. Climate data included 30-year (1961–1990) average monthly and annual precipitation and temperature measures, represented by spatial resolutions of about 4 km for the conterminous U.S. (Daly et al., 2000), 2 km for Alaska (Daly, 2002) and 420-m for Puerto Rico (Daly et al., 2003). The dataset also included elevation from 30-m digital elevation models (DEMs) (Gesch et al., 2002), and other topographic derivatives from those DEMs, including slope, dominant aspect, and an indicator of aspect variety. This indicator is calculated as the total number of unique aspect values (or the variety) within the nine by nine window surrounding each 30-m cell. The resulting dataset was resampled to a 250-m cell size. The same resampling method used for the 30-m Landsat products (described above) was used to summarize the elevation-based attributes at 250-m. A final topographic variable that several models used was a horizontal-distance-to-nearest-stream measure, which is the Euclidean distance from each pixel to its nearest above-ground water body, as the crow flies.

2.2. Modeling strategy

2.2.1. Process overview

We created a nationwide modeling dataset by intersecting plot locations with the geospatial predictor layers, and extracting all relevant data. Resulting values of predictor layers for each plot were then linked to the corresponding forest/nonforest and forest biomass response variables. We segmented this modeling dataset into 65 ecologically unique mapping/modeling zones (Fig. 1) (Homer & Gallant, 2001) which permitted separate models to target the conditions unique to each zone. However, we aggregated adjacent zones in sparsely forested regions, which had too few forested plots, to increase the number of observations in the models for those zones. Independent test sets were created by randomly selecting 10 to 15% of the plots by mapping zone, leading to proportional distribution by zone. These test sets were withheld to assess model performance, except in Puerto Rico where insufficient numbers of plots forced the use of 10-fold crossvalidation for evaluating biomass model performance. Using classification trees with boosting for each mapping zone, we first produced a 250-m resolution forest mask by modeling the binary variable of forest/nonforest as a function of all the variables contained in the national geospatial predictor layers. We then selected only those FIA plots that fell within the forested portion of the forest mask as training data for the biomass models. Regression tree algorithms were then used to model forest biomass (also at 250 m) as a function of those same predictor variables used in the forest/nonforest models for each mapping zone. Because of ecological differences between zones, the way in which the classification trees used and partitioned the predictor variables was very different by zone. Also, some regional variations in the methods themselves were used to improve the forest/nonforest and biomass models. Examples include inclusion of regional specific predictor layers and larger groupings of similar mapping zones. We then predicted forest biomass on a per-pixel basis by applying the models developed for each mapping zone to the corresponding predictor layers for that zone. Pixels with nonforest class label predictions were omitted from subsequent analyses, and labeled as having no forest biomass. Finally, using the classification confidence and absolute error information available from the models, two additional geospatial datasets were created to capture the per-pixel uncertainty associated with each estimate – resulting in a map of forest probability, and a map of biomass percent error (details in section on Uncertainty maps). The individual zone maps of forest/nonforest, forest probability, biomass, and percent error for biomass were mosaiced to form nationwide datasets. A state boundary geospatial layer identified coastal shorelines (nationalatlas.gov/statesm.html), and a national hydrography layer (nationalatlas.gov/hydrom.html) delineated interior water boundaries.

2.2.2. Classification and regression trees

Classification and regression tree modeling, or recursive partitioning regression (Breiman et al., 1984), is available in many software packages and is now common in remote sensing applications. To give a general overview of the methodology, trees subdivide the space spanned by the predictor variables into regions for which the values of the response variable are most similar, and then assign a unique prediction for each of these regions. The tree is called a classification tree if the response variable is discrete and a regression tree if the response variable is

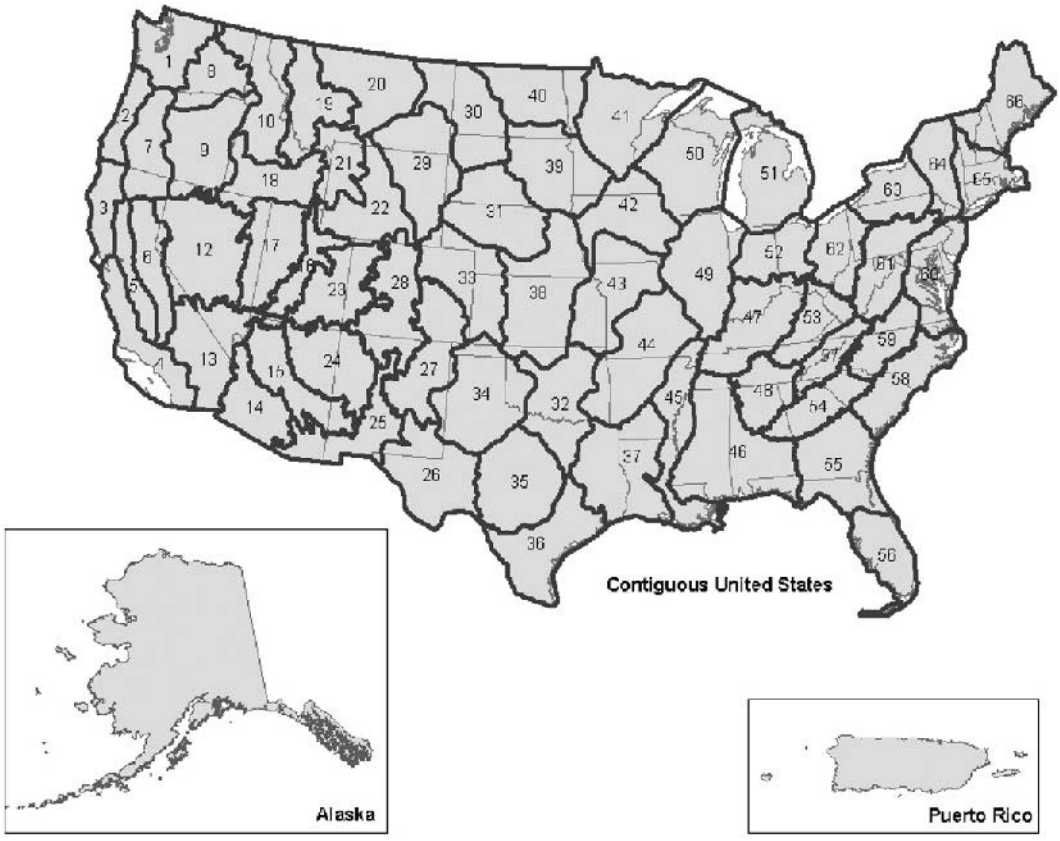

Fig. 1. Mapping/modeling zones (Homer & Gallant, 2001) segmented forest vs. nonforest and biomass classification and mapping models.

continuous. Tree-based methods have evolved to enhance their predictive capabilities. Two recent enhancements have had considerable success in mapping applications (Chan et al., 2001). One is known as bagging, or bootstrap aggregation (Bauer & Kohavi, 1998; Breiman, 1996). The other is called boosting (Freund & Schapire, 1996) with its variant Resampling and Combining (ARCing) (Breiman, 1998). These iterative themes each produce a committee of expert trees by resampling with replacement from the initial data set, then averaging the trees with a plurality voting scheme if the response is discrete, or simple averaging if the response is continuous. The difference between bagging and boosting is the type of data resampling. In bagging, all observations have equal probability of entering the next bootstrap sample. In boosting, problematic observations, those which are frequently misclassified, have a higher probability of selection. The performance of tree-based methods for modeling FIA response variables is compared to other modeling techniques in Moisen and Frescino (2002) and Moisen et al. (2006).

Specifically for this study, classification trees with boosting (5 or 10 trials) and pruning in See5 (www.rulequest.com, Quinlan, 1986, 1993) generated the forest mask based on a 0.5 threshold for distinguishing forest from nonforest. Cubist (www.rulequest.com) generated the mapping models of forest biomass within pixels predicted to be forested. Cubist is a proprietary variant on regression trees with piecewise nonoverlapping regression. Specific software options used for most mapping zones included the following: either 5 or 10 committee models; use of rules alone (no instances); minimum rule cover of 1% of cases; extrapolation up to 10%; and no maximum number of rules.

2.3. Model performance

Measures for assessing and depicting accuracies, errors, and uncertainties of the modeled spatial datasets were chosen by taking into consideration traditional methods of accuracy assessment, known characteristics of the datasets, and their anticipated uses.

2.3.1. Per-pixel measures

Accuracy and error measures for the forest mask included proportion of correctly classified units (PCC), Kappa (Cohen, 1960), as well as omission and commission errors for both the forest and nonforest classes. PCC is a statistic that can be deceptively high when the proportion of a class, in this case forest, is very low or very high. The Kappa statistic measures the proportion of correctly classified units after removing the probability of chance agreement. Errors of omission (1-producer's accuracy) result when a pixel is incorrectly classified into another category, thus being omitted from its correct class. Errors of commission (1-user's accuracy) result when a pixel is committed to an incorrect class. For the biomass map, the per-pixel accuracy measures that we calculated on the independent test sets included average absolute error,

relative error, and correlation. The average absolute error for a set of test cases is the average of the sizes of differences between the actual and predicted values for each case, expressed in metric tons per ha. The relative error is the ratio of the average absolute error to the average absolute error that would result by predicting the value of each case as the mean of the training set. Because it is normalized by the predicted value's unit of measure, the relative error term is useful for comparing the performance of different models. It also gives an indication of individual model performance above and beyond simply using the average value from the training data as its 'predicted' value. A relative error substantially less than one indicates that the model predictions are substantially better than simply using a prediction of the sample mean. The correlation coefficient is a standard measure of the linear relationship between observed and predicted values.

2.3.2. Uncertainty maps

One of the goals of this study was to provide spatially explicit depictions of the uncertainty in both the forest mask and forest biomass maps. Maps of uncertainty are derived from the modeling process itself and provide users (and developers) information on where the model was more and less confident of the estimate based on the training and predictor information available and the modeling technique used.

For the forest/nonforest map, a binary response variable, the need for a spatial depiction of uncertainty was satisfied with a forest probability dataset, depicting the probability that any individual pixel could be classified as forest. In many modeling applications for binary response variables, predictions are made on a continuum of 0 to 1, indicating probability of a pixel belonging to the class of interest. Because of the way in which See5 constructs predictions, a map of forest probability had to be back-engineered in the following way. First, the public C code distributed with See5 (http://rulequest.com/see5-public.zip) enabled us to produce a confidence value for each pixel prediction as a forest/nonforest classification confidence map. This software routine operates as follows: if a single classification tree is used and a case is classified by a single leaf of a decision tree, the confidence value assigned is the proportion of training cases at that terminal node that belongs to the predicted class. If more than one terminal node is involved, the confidence value assigned is a weighted sum of the individual nodes' confidences. If more than one tree is involved (eg. boosting), the value is a weighted sum of the individual trees' confidences. Second, a forest probability map was created by remapping confidence values from the public C code to a range of 0 to 0.5 for nonforest pixels and 0.5 to 1 for forest pixels, creating a new range from 0 to 1. Here, values near 0 indicate a more confident prediction for nonforest areas, values near 1.0 indicate a more confident prediction for forest areas, and values around 0.55 are the most uncertain.

For the map of aboveground forest biomass, spatial depictions of uncertainty took the form of biomass percent error maps. These were derived by first extracting the weighted average absolute error of all the rules that applied to each pixel, in which the average absolute error for each rule is from the training data. The biomass percent error map then resulted from dividing that weighted average absolute error by the predicted biomass value at that pixel. Such uncertainty maps provide information regarding both the location and magnitude of potential errors in the modeled estimates. They allow users to incorporate this information into all further modeling or analysis efforts using the estimated biomass and forestland maps/datasets (Fortin et al., 1999; Mowrer, 1994; Woodbury et al., 1998).

2.3.3. Agreement of spatial aggregations

FIA plot data is typically used to produce unbiased estimates of forest population totals using design-based inference (Cochran, 1977; Särdnal et al., 1992; Thompson, 1997) for areas of sufficient size. Often in practice, however, maps may be used to produce population estimates of these mapped variables by summing pixels over the geographic area of interest. This method relies on model-based inference (Valliant et al., 2001). To provide information on the comparative accuracy of these "map-based" estimates of area of forestland and total biomass, we compared them to "plot-based" estimates of total forest area and biomass by state for the US using FIA sample plots (Scott et al, 2005). Note that although FIA will use remote sensing information to stratify sample plots to improve precision in estimates of forest population totals, the plot-based estimates used here are solely based on field data. This comparison allows users of inventory data who are familiar with the traditional plot-based estimates to examine the location and magnitude of areas of over-and underestimation of map-based estimates.

Next, in order to examine the scales at which aggregated estimates of forest area or total forest biomass agree with plot-based estimates, we also made comparisons for hexagons at four different sizes: ~16,000, ~21,000, ~39,000, and ~65,000 ha. The hexagons were derived by tessellation from the Environmental Monitoring and Assessment Program hexagons (White et al., 1992) that are used as the basis for the FIA sampling design (Bechtold & Patterson, 2005). For both area of forestland and

Table 1
Per-pixel measures of performance for forest/nonforest maps based on independent test sets, reported by region

Region	PCC	Kappa	Omission forest	Commission forest	Omission nonforest	Commission nonforest	Test set sample size
Northeast	0.89	0.77	0.08	0.09	0.14	0.14	1181
Northcentral	0.93	0.80	0.15	0.15	0.05	0.05	5449
Interior West	0.91	0.76	0.17	0.18	0.07	0.06	7196
Pacific Northwest	0.85	0.61	0.05	0.15	0.39	0.15	2588
Southern	0.86	0.69	0.10	0.13	0.22	0.17	3138
Alaska	0.94	0.88	0.07	0.08	0.05	0.04	6553
Puerto Rico	0.79	0.57	0.07	0.28	0.36	0.10	28

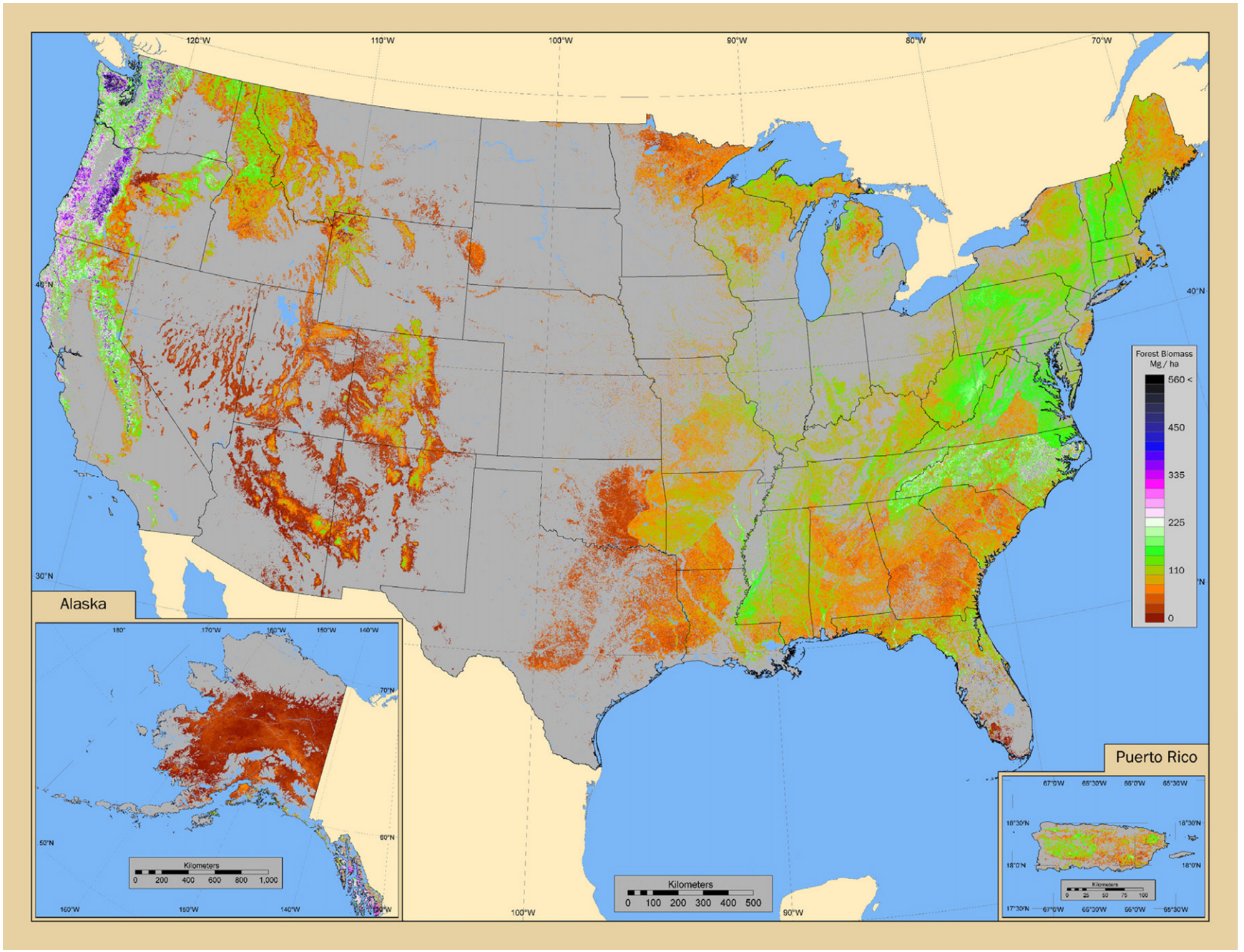

Fig. 2. A map of aboveground forest biomass (dry weight) in live trees, tumps, branches and twigs derived from modeling FIA plot biomass as a function of geospatial predictor variables.

Table 2
Per-pixel measures of performance for biomass maps based on independent test sets (except for Puerto Rico where 10-fold cross-validation was used), reported by region

Region	Average absolute error	Relative error	Correlation	Test set sample size
Northeast	60.1	0.89	0.39	1156
Northcentral	42.5	0.88	0.46	1134
Interior West	42.2	0.65	0.66	2023
Pacific Northwest	163.1	0.72	0.73	1591
Southern	60.2	0.92	0.31	1939
Alaska	91.5	0.59	0.69	430
Puerto Rico	65.0	0.51	0.92	*

*Based on a 10-fold cross validation.
Average absolute error is reported in metric tons per hectare.

aboveground forest biomass, agreement between the mean of pixel predictions for all pixels with centers in a hexagon to the mean of plot observations for all plots with centers in the hexagon was assessed as follows. For each hexagon, the mean pixel prediction, $\hat{\mu}_{\text{pixel}}$, for a hexagon was compared to the plot-based mean, $\hat{\mu}_{\text{plot}}$, using,

$$\tau = \frac{\hat{\mu}_{\text{pixel}} - \hat{\mu}_{\text{plot}}}{\text{SE}(\hat{\mu}_{\text{plot}})}$$

where $\text{SE}(\hat{\mu}_{\text{plot}})$ denotes the design-based standard error. Here, τ is not a formal statistic with an established distribution and probability levels. Rather it is constructed as a heuristic tool by with which to assess relative agreement between traditional plot-based estimates, and map-based estimates at varying scales of aggregation.

Arrays of hexagons of four sizes were considered: ~16,000 ha, ~21,000 ha, ~39,000 ha, and ~65,000 ha. Based on a sampling intensity of approximately one plot per 2400 ha, hexagons of ~16,000 ha would include 6–7 plots, about the smallest sample sizes that would yield reliable estimates of SE ($\hat{\mu}_{\text{plot}}$). Selection of the sizes of the larger hexagons was arbitrary, except for the ~65,000-ha EMAP hexagons, from which unique ID codes are attributed to FIA plots and which are used in several national assessments. For areas of the country in which a complete cycle of sampling has not been completed, some hexagons may include fewer than 6–7 plots. No comparisons of pixel-and plot-based means were calculated for hexagons with fewer than 5 plots. Tau-values exceeding 2 were interpreted as a conservative indication that model-based estimates disagreed with plot-based estimates within each hexagon.

2.3.4. Variable importance

A series of variable importance analyses were conducted to assess the relative contributions of the numerous predictor variables to the modeling process. First, the relative importance of the major groups of predictor variables were assessed in each region. This was measured by percent improvement, or decrease, in relative error when each major group was used alone as predictor sets in different models of biomass. Major groups were the "MODIS group" (including NDVI, Enhanced Vegetation Index [EVI], spectral bands, fire, and percent tree cover), the "Climate group" (including all precipitation variables), the "NLCD group" (including only NLCD-derived variables), and the "Topo group" (including topographic variables). Note that in Alaska and in Puerto Rico, no NLCD data were available, and surrogate variables labeled as the "Veg group" were used instead.

Next, the relative importance of sub-groups of the "MODIS group" were measured by percent improvement, or decrease, in relative error when each of these sub-groups was used exclusively in the models. These sub-groups were the "Bands sub-group" (including all MODIS bands, all dates), the "NDVI sub-group" (including al NDVI variables), "Treecov sub-group" (including percent tree cover), the "EVI sub-group" (including all EVI variables), and the "Fire sub-group" (including all fire-related variables).

Because the true contribution of each variable to the final biomass map is confounded by high correlation between variables, variable groups were excluded in turn from the original biomass model, and the effect on relative error examined. In addition, the potential effect on pixel aggregations were explored by examining changes in density functions of predicted values under the different models excluding variable groups in turn.

2.3.5. Estimates of C pools

Finally, estimates of C pools in live forest biomass of U.S. forests, derived from the map developed in this study, were compared with estimates from other national studies. Estimates of the mass of C for live trees, stumps, branches and twigs were obtained by summing one-half the predicted biomass for each pixel over the conterminous U.S., and Alaska. The one-half rule is based on Brown and Lugo (1992). Mass of C for roots was approximated as 20% of total predicted biomass (Cairns et al., 1997). Results were compared those obtained by Turner et al. (1995), Birdsey and Heath (1995), Potter (1999), and Dong et al. (2003).

3. Results

All maps produced in this study, including the forest/nonforest mask, forest probability, forest biomass, and biomass percent error, are available for download via http://svinetfc4.fs.fed.us/rastergateway/biomass/.

3.1. Per pixel measures

As illustrated in Table 1, the forest mask was reasonably accurate in all regions, with regional PCCs ranging from 0.79 in Puerto Rico to 0.94 in Alaska, and regional Kappa values ranged from 0.57 in Puerto Rico to 0.88 in Alaska, reflecting fair to excellent class agreement. Errors of omission for forest were generally low, ranging from 0.05 in the heavily forested Pacific Northwest to 0.17 in the more nonforested Interior West, while errors of commission for forest ranged from 0.08 in Alaska to 0.28 in Puerto Rico. Errors of omission for nonforest ranged from 0.05 in the Northcentral region and

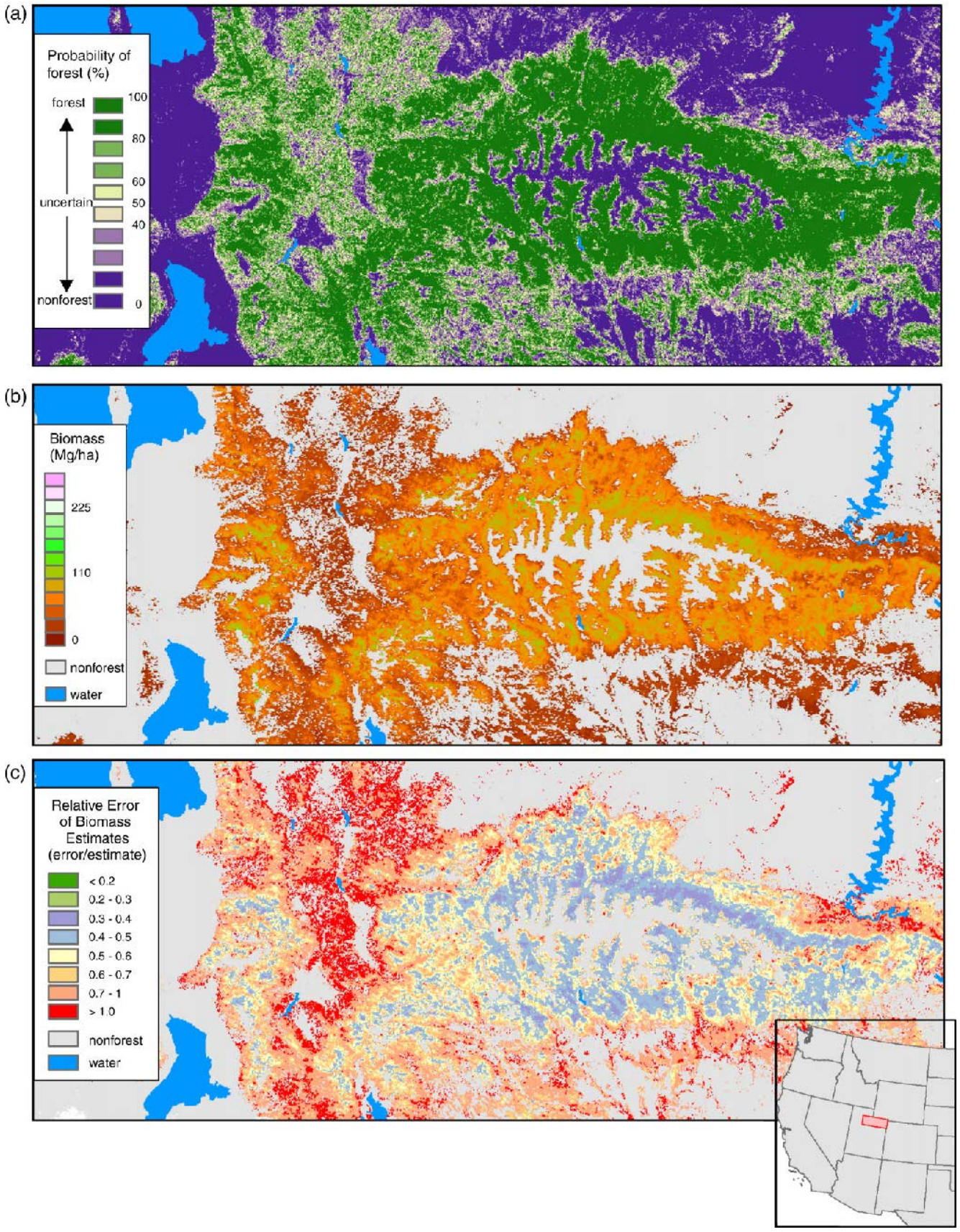

Fig. 3. Probability of forest, biomass, and percent error in biomass mapped over Uinta Mountains in Utah, (a, b, and c respectively). Probability of forest, biomass, and percent error in biomass mapped for the Greater Mohawk Valley Region, New York (d, e, and f respectively).

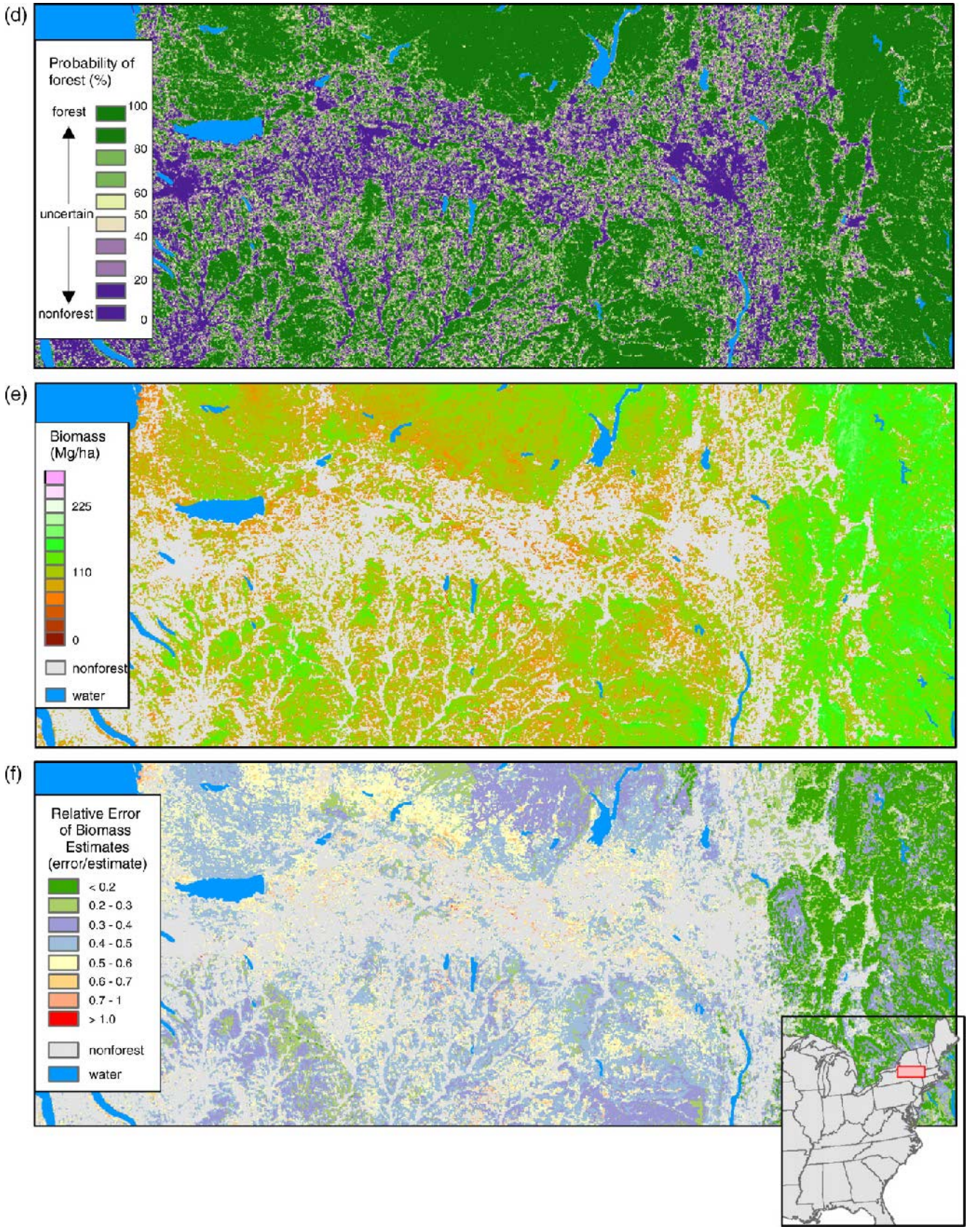

Fig. 3 (*continued*).

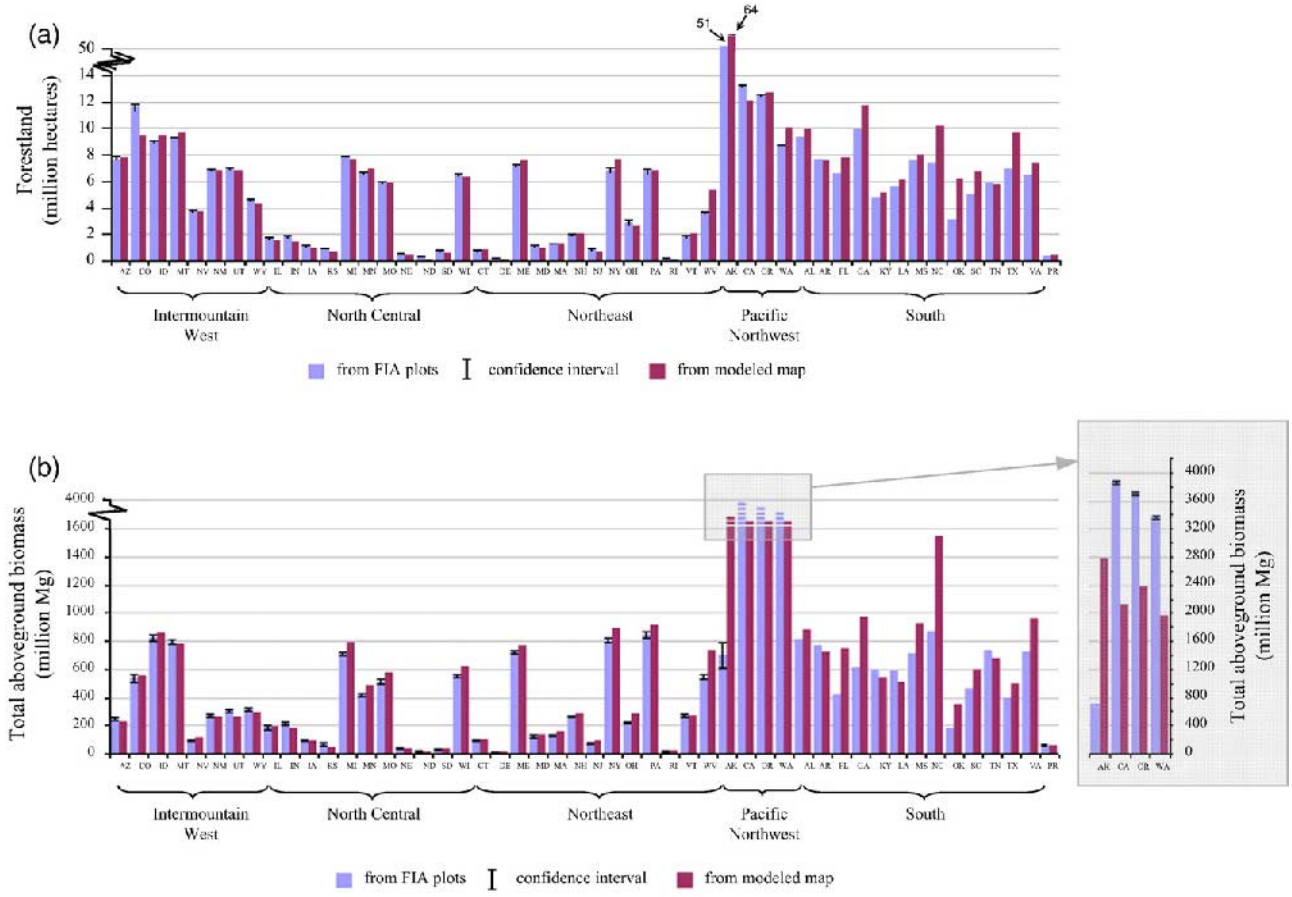

Fig. 4. Plot-based and map-based estimates of (a) forestland and (b) total forest biomass (dry weight), by state. States are grouped by USFS Forest Inventory and Analysis Region. A separate *Y* axis is provide for the Pacific Northwest states because of the substantially different scales involved.

Alaska to 0.39 in the Pacific Northwest, while errors of commission for nonforest ranged from 0.04 in Alaska to 0.17 in the Southern region. Per-pixel measures of performance for the forest/nonforest maps are given for individual and aggregated zones in Appendix A.

The forest biomass map is presented in Fig. 2. The models of aboveground live forest biomass varied by region in their ability to predict pixel-level values (Table 2). Correlation coefficients ranged from 0.92 in Puerto Rico down to 0.31 in the Southern region. The western regions had substantially better results than did those in the eastern regions of the US. Relative errors ranged from 0.51 in Puerto Rico to 0.92 in the Southern Region, with the former value indicating an approximate 50% improvement over using the sample mean from the model's training dataset, versus a more modest improvement in performance over a simple sample mean indicated by the latter value. Most individual mapping zones (75%) had relative errors less than 1.0, indicating gains in the modeling process. However, some zones actually had a relative error greater than 1.0 indicating the models performed worse than using a simple sample mean. This was particularly true in zones with a high proportion of scattered forest that is hard to identify with a 250 m pixel (e.g., zones 52, 44, and 49) and/or

Table 3
Assessment of agreement between plot-and map-based estimates of forest land area and total biomass over 4 scales of spatial aggregation across the continental US

Hexagon size (ha)	Estimate	Number of hexagons	Average plots/ hexagon	Proportion of hexagons				
				$-3<\tau$	$-3\le\tau<-2$	$-2\le\tau<2$	$2\le\tau<3$	$\tau>3$
16,000	Forest area	25,512	9.40	0.013	0.011	0.938	0.026	0.012
21,000	Forest area	22,327	10.73	0.014	0.012	0.931	0.030	0.014
39,000	Forest area	15,993	14.99	0.019	0.019	0.908	0.036	0.018
65,000	Forest area	10,439	22.96	0.023	0.026	0.879	0.047	0.026
16,000	Biomass	25,512	9.40	0.003	0.007	0.887	0.042	0.061
21,000	Biomass	22,327	10.73	0.004	0.008	0.879	0.046	0.063
39,000	Biomass	15,993	14.99	0.005	0.012	0.860	0.049	0.074
65,000	Biomass	10,439	22.96	0.008	0.019	0.835	0.051	0.087

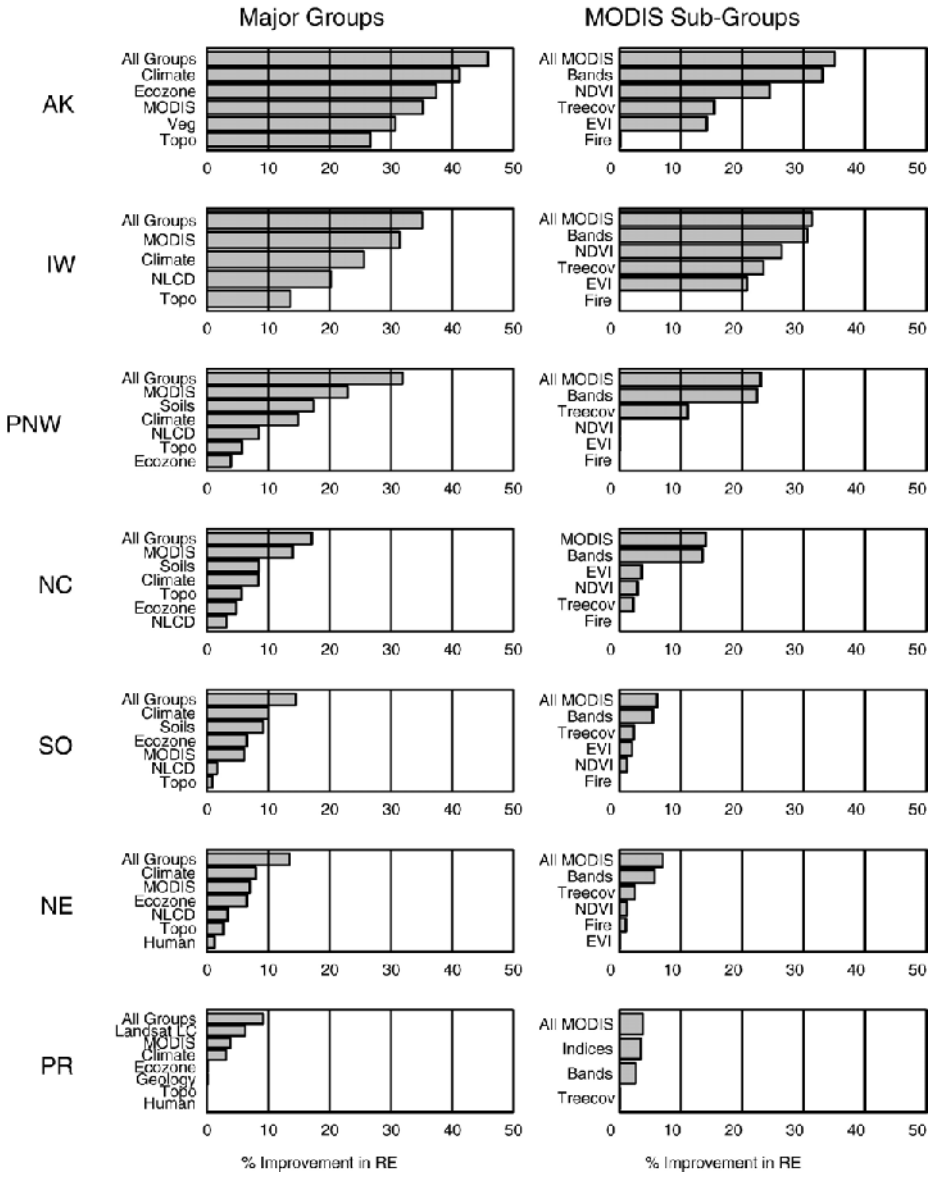

Fig. 5. Relative importance of the major groups of predictor variables as well as sub-groups of MODIS variables in each region. Importance is measured as the percent improvement in relative error when each variable group is used individually in a model of forest biomass. Regional abbreviations include: AK - Alaska, IW - Interior West, PNW - Pacific Northwest, NC - North Central, SO - Southern, NE - Northeast, and PR - Puerto Rico.

areas missing forest data (e.g., zones 32 and 35). Biomass model performance results are given for individual and aggregated zones in Appendix A.

3.2. Uncertainty maps

The forest probability map reflects uncertainty in pixel assignments to forest or nonforest categories in the forest mask. The forest probability map is a useful product of the forest-nonforest modeling process because it allows users to choose their own application-specific threshold for distinguishing between forested and nonforest lands. The biomass percent error map reflects uncertainty in the modeled pixel-level biomass values.

In general, the uncertainty maps reflect those areas that are more difficult to model because of their spatial characteristics, because of poor quality training or predictor data available in those areas, or because of a poor relationship between the desired response variable and the predictor layers available. In the forest probability map these were the interface areas between forest and nonforest, and in the forest biomass map these were the areas that were less intensely sampled, more affected by land use history (which was not an available predictor layer) or otherwise difficult to model.

Looking more closely at the resulting biomass, biomass uncertainty, and forest probability maps, regional differences in patterns of map uncertainty are apparent. In Fig. 3a, the Uinta Mountains of the Interior West, large areas of highly certain

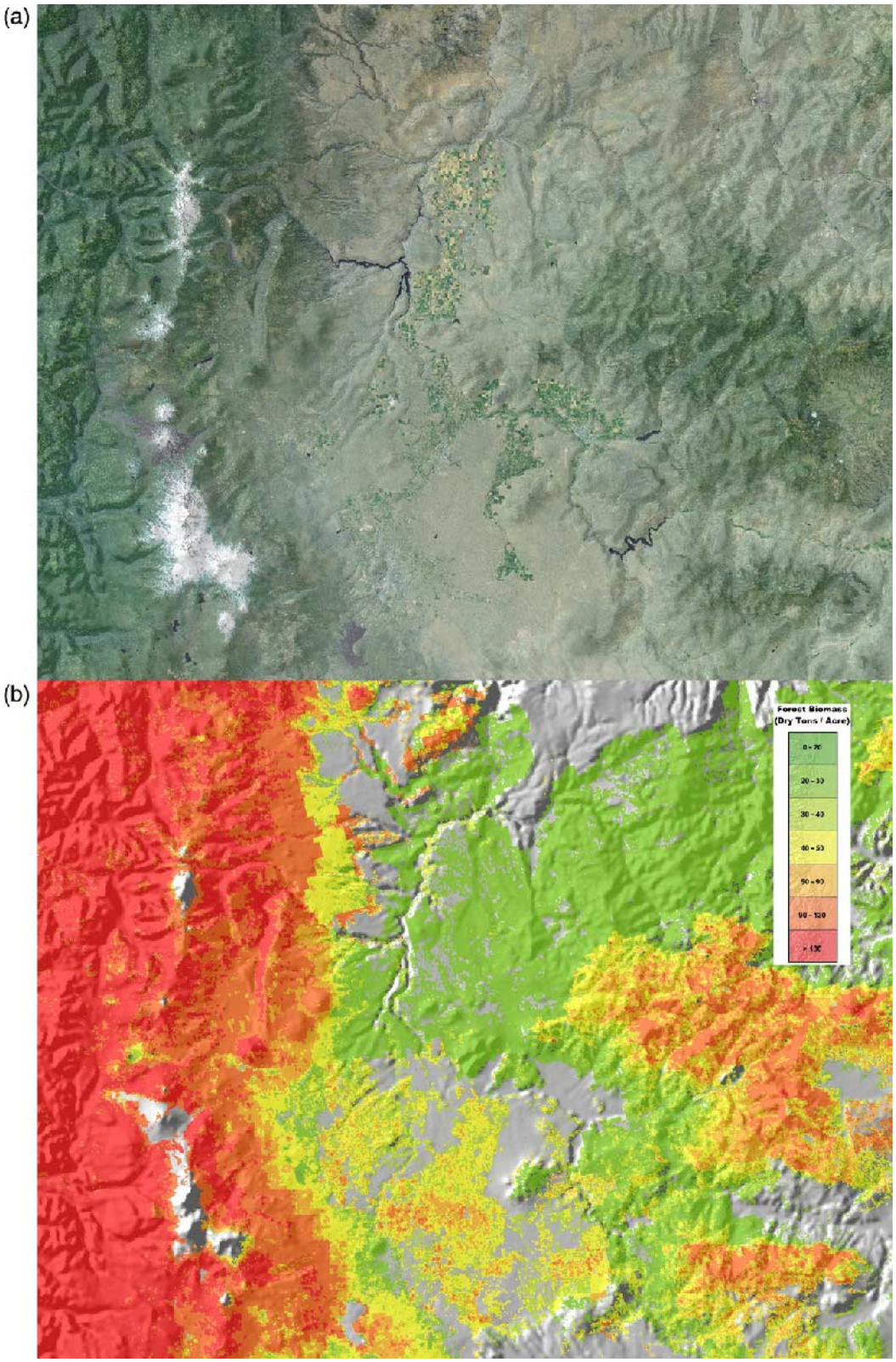

Fig. 6. An enlarged view of the natural color MODIS imagery (a), and the corresponding biomass dataset (b) from the Pacific Northwest in central Oregon.

nonforest exist relatively unbroken by pixels with much probability of forestland. In contrast, in the greater Mohawk Valley region of the Northeast (Fig. 3d), there are few continuous areas of highly certain nonforest.

In both regions it was the spatially heterogeneous areas that were the most difficult to predict — the highly intermixed forest-agriculture and forest-developed interfaces in the Northeast, and the sparse canopy transition zones between forest and nonforest

Table 4
Effect of excluding variable groups on relative error by region

Variable groups excluded	Increase in relative error						
	NE	NC	SO	INT	PNW	AK	PR
MODIS	0.01	0.04	0.03	0.02	0.05	0.03	0.02
Topo	0.00	0.00	0.01	0.01	0.01	0.02	−0.03
NLCD/Veg†	0.00	0.00	0.01	0.01	0.02	0.01	0.02
Climate	0.00	0.00	0.01	0.00	0.01	0.03	0.00
Soils/geology	–	0.01	0.01	–	0.02	–	0.00
Ecozone	0.00	0.01	0.01	–	0.01	0.01	−0.02
Human	0.00	–	–	–	–	–	0.00

†Land cover data for Puerto Rico from Helmer et al. (2002).
Increase in relative error is measured as the difference between relative error obtained excluding each of the major predictor groups in turn, and the relative error obtained using all the predictor variables.

areas at the elevational (i.e. treeline) and arid limits of tree growth in the Interior West. In both regions, the probability of forest values falling in the most uncertain range (0.4 to 0.6) represented just over 10% of the dataset — a substantial portion, illustrating the difficulty of accurately determining this edge, particularly at this resolution.

The uncertainties associated with biomass predictions in the Interior West are strongly related to the amount of biomass present, with higher percent errors associated with the lower biomass values (Fig. 3b). In the Northeast, percent errors were lower in general and show a spatial pattern that differs from the biomass predictions themselves (Fig. 3e). This pattern may reflect the distribution of different types of forest and our ability to model biomass in each, but also likely is the influence of the ancillary layers used in the modeling. Without the strong influence of a single variable, such as elevation, biomass predictions in the Northeast relied upon different predictor layers in different areas, each with varying levels of confidence that seemed to be visually correlated with these layers. Percent error values were in general much higher (above 0.8) in the Interior West (Fig. 3c) than in the Northeast (Fig. 3f). This is in large part due to the relatively lower biomass values present in the Interior West as compared to the Northeast.

3.3. Agreement of spatial aggregations

As described in Section 2.3.3, estimates of total forest area and biomass were computed by state from FIA sample plots. These plot-based estimates were compared to map-based estimates of total forest area and biomass that resulted from counting forested pixels and summing their biomass. At the state level, spatial aggregation results show fairly good agreement between the two sources for forest area, with notable exceptions in CO (where the map underestimates forest area) and GA, WV, NC, TX, and OK (where the map overestimates forest area). Twenty-nine of the states' map-based estimates fell within 10% of the plot-based estimates for forest area (Fig. 4a). For aboveground forest biomass, spatial aggregation results show an overestimation of biomass in most areas, with the notable exceptions of CA, OR, and WA where the map appears to substantially underestimate forest biomass. Substantial overestimation of state-level summaries appeared to occur in NC, VA, GA, AK, CA, OR, WA, and WV. Twenty one of the states' map-based estimates fell within 10% of the plot-based estimates for biomass (Fig. 4b).

Table 3 illustrates the distribution of τ-values to assess agreement between plot-and map-based estimates of total forest area and total biomass at four spatial scales of aggregation across the continental US. Map-based estimates of forest area generally were in agreement with plot-based estimates for all hexagon scales. However, spatial aggregations of hexagons with large absolute τ-values indicate that the forest mask is problematic in some portions of the Southeast; probably in parts of Maine, Wisconsin, Minnesota, Oklahoma, and along the Great Lakes; and perhaps in parts of the Pacific Coast states. These aggregations of hexagons with disagreeing estimates appear to be consistent across all four hexagon scales. Not surprisingly, the biomass map appears to exhibit more disagreement than observed for the forest mask at each hexagon aggregation. However, most disagreement in the biomass map resulted from over-estimates, while disagreement in the forest mask appeared more evenly distributed between over-estimates and under-estimates.

3.4. Variable importance

The first column in Fig. 5 depicts the relative importance of the major groups of predictor variables in each region. This was measured by percent improvement, or decrease, in relative error when each of these variable groups was used alone as predictor sets in different models of biomass. The bar labeled "All Groups" illustrates the maximum decrease in relative error obtained by including all the predictor variables, indicating improvement over just using the sample mean. In four of the seven regions (NC, IW, PNW, and PR), the "MODIS group" resulted in the largest improvement in relative error. In the other three regions (NE, SO,

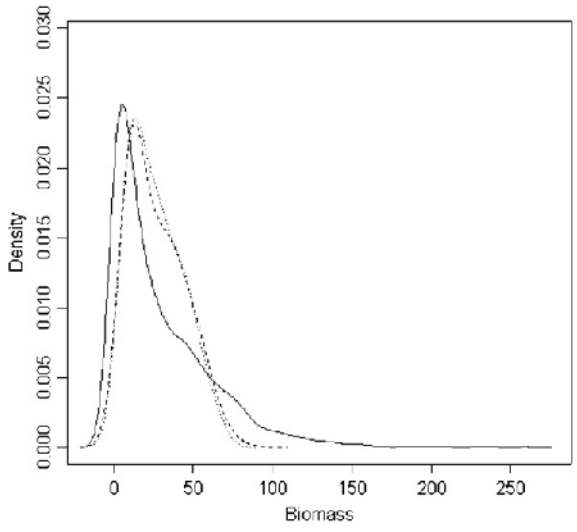

Fig. 7. The density function in the Interior West of observed biomass values (solid line), as compared to that from a model containing all the predictor variables (dashed line), and from a model excluding all the MODIS-derived variables (dotted line).

Table 5
Estimates of C pools in live forest biomass of continental U.S. forests

Source	Approach, spatial resolution, and study area addressed	Forest components	Forest biomass estimates for the U.S. (Pg C)[†]
Not spatially explicit			
Turner et al. (1995)	Inventory data by forest type at State level (1980–1990) for conterminous U.S.	Live trees, stumps, roots, branches, twigs and shrubs	14.96[‡]
Birdsey and Heath (1995)	Inventory data by forest type at State level (1980–1992) for continental U.S.	Live trees, stumps, roots, branches, twigs, shrubs and herbs	16.74
Spatially explicit			
Potter (1999)	Satellite-image scaled physiological model at 1° (1980s — ignores forest age structure) for the Earth	Live trees, stumps, roots, branches, twigs, and leaves.	37.65
Dong et al. (2003)	Inventory data at Province level scaled with satellite imagery to 8 km (1990–1995) for Northern Hemisphere temperate and boreal countries	Live trees, stumps, roots, branches, twigs and shrubs	12.48
This study	Inventory data at plot level scaled with satellite imagery to 250 m (2001) for continental U.S.[†]	Live trees, stumps, roots, branches and twigs	18.08[§]

[†]All estimates exclude Hawaii and Puerto Rico. This study estimates that Puerto Rican forests have 53.4 Mg C in aboveground live forest biomass.
[‡]Including 12.6 Pg C for conterminous U.S. plus 2.36 Pg C for Alaska from Birdsey and Heath (1995). [§]Includes root biomass estimated as 20% of total biomass (Cairns et al., 1997).

and AK) the "Climate group" resulted in the largest improvement in relative error. Use of just the "NLCD" and "Topo" groups alone resulted in smaller improvements in relative error than the "MODIS" or "Climate" groups in all regions. Some regions also opted to include additional variables groups related to soils, development, etc. Although not common to all regions, they are shown here for comparison sake. Note that because of high correlation between variables, the sum of decreases in relative error realized by modeling the groups individually cannot be expected to equal the total decrease in error when modeling all variables together. Variable groups contain redundant information, as will be illustrated later.

The second column in Fig. 5 depicts the relative importance of sub-groups of the MODIS-based variables as measured by percent improvement, or decrease, in relative error when each of these variable groups is used exclusively in the models. The bar labeled "All MODIS" provides a reference for the maximum decrease in relative error possible by using all the MODIS variables together. Using the "Bands" group alone (including all MODIS bands, all dates) resulted in models that performed nearly as well in most regions. Use of just the "NDVI" variables, "Treecov" (percent tree cover) variable, and "EVI" variables resulted in progressively smaller decreases in relative error. Note that the fire-related variables made no contribution when used alone. As with the major groups, sub-groups of variables within the MODIS group contain redundant information resulting in non-additivity of their relative contributions. Fig. 6a is an enlarged view of the natural color MODIS imagery, and the corresponding biomass dataset is shown in Fig. 6b. The image is from the Pacific Northwest in central Oregon, which visually demonstrates the high degree of correspondence between the MODIS data and biomass predictions.

While the results above illustrate the relative predictive information contained in each groups or sub-groups of variables, the true contribution of each variable to the final biomass map is confounded by high correlation between variables. Consequently, variable groups were excluded in turn from the original biomass model, and the effect on relative error shown in Table 4. In all cases except the MODIS group, exclusion of these variables resulted in a 2% or less change in relative error. Exclusion of the MODIS group had the largest impact over the other groups in all regions, although that impact, too, was very small, ranging from only 1% in the NE to 5% in the PNW. Also noted was the negative, albeit small, impact of including groups of variables exhibiting no contribution to the biomass prediction in Puerto Rico, where small sample size made models more vulnerable to extraneous information.

Not only was there minimal effect on pixel-level accuracies, the potential effect on pixel aggregations can be surmised by examining changes in density functions of predicted values under different models. Fig. 7 illustrates the density function in the Interior West of observed biomass values (solid line), as compared to that from a model containing all the predictor variables (dashed line), and from a model excluding all the MODIS-derived variables (dotted line). Both the all-variable model and the model excluding all the MODIS variables result in nearly identical densities. This illustrates the tendency in all these models to predict closely to the mean and not capture the observed variability in biomass. We also observed a very large discrepancy between variances for observed and variances for predicted values. As a side note, a likely contributor to this phenomenon is the spatial resolution (pixel size) at which the models are implemented. We only have biomass observations from small field plots and are modeling these to biomass on 250-m pixels. Yet we know that as pixel size increases, pixel values become more like the mean, and variance decreases. This will be addressed further in the discussion. But the differences between predicted value variances resulting from models excluding different groups of variables in turn are quite small. Because of redundancy of information between predictors, exclusion of any one of the major groups had only a small effect on the prediction accuracies and aggregations.

3.5. Estimates of C pools

Carbon pool estimates in live forest biomass of U.S. forests, derived from the map produced in this study, compare well with estimates from other studies (Table 5). The estimates for U.S. forests from Turner et al. (1995) and Birdsey and Heath (1995) are strictly plot-based (with the exception of Alaska), and they use FIA data from the 1980's to early 1990's. The estimate from Potter (1999) is from a global study and is high because it ignores forest age structure. It scales AVHRR NDVI data with a biophysical model, estimating potential forest biomass of forested areas. Dong et al. (2003) address temperate and boreal forests of the Northern Hemisphere. They scale state-and province-level estimates of total forest biomass from forest inventory data with cumulative NDVI indices from AVHRR data. It is the smallest estimate for around 1990 and may indicate that using satellite imagery to scale state-level forest biomass underestimates forest biomass. These forest carbon estimates probably also differ because the scales of these studies range from national to global.

4. Discussion

Image products from MODIS were useful for this study not only because they were practical, but also because they were preferable for scaling reasons. From a practical standpoint, the coarser spatial resolution of MODIS imagery makes applications at sub-continental scales computationally less intensive compared with finer resolution data. Moreover, MODIS image products, like tree cover data and preprocessed image composites that minimize cloud cover, along with the larger scene and tile sizes, reduce the burdens of image preprocessing. At the same time, the land imaging MODIS bands include optical bands comparable to finer scale data. These bands center on visible, near infrared and shortwave infrared bands that many studies show are sensitive to forest cover and, within limits, forest stand structure. Bands 1 and 2 of MODIS, for instance, are centered on the red and near infrared parts of the electromagnetic spectrum and are important in indices sensitive to photosynthetic vegetation. Bands 2 and 6 are similar to Landsat image bands 4 (near infrared) and 5 (shortwave infrared), respectively, which form indices sensitive to forest structure or successional stage in both temperate (Fiorella & Ripple, 1993) and tropical (Helmer et al., 2000) landscapes.

From a scaling perspective, the 250 to 500-m pixel size of MODIS bands 1–2 and 3–7, respectively, were beneficial overall. Variable importance analyses revealed that MODIS-derived information could contribute more predictive power than other classes of information when used in isolation. However, because of strong correlation between variables, the true contribution of MODIS-derived variables when used in concert with the broad suite of other predictors was quite small. In addition, the coarse scale likely added to plot-pixel differences. A summary of possible sources of per-pixel errors in the biomass map would include: 1) reflectance values in dense canopy forests saturate at relatively low levels of forest biomass, 2) the spatial mismatch between the FIA plots and the 250-m pixels, and 3) errors in the forest/nonforest mask. With 250-m pixels, positional inaccuracy is unlikely to contribute to model errors, though it could be a factor.

First we address the saturation of reflectance values. In most mapping zones, the models tended to underestimate large biomass densities and overestimate small ones, truncating the range of values predicted and adding to the average relative error in models. The most important source of these residual errors in mapping models probably stems from the well-known fact that canopy reflectance from passive optical sensors has limited sensitivity to the canopy structure of dense forests, where most live forest biomass is. Forests continue to accumulate biomass after canopies close as well as after indices of vegetation greenness and net primary production level off. Yet this very limitation was one of the reasons why we worked at a spatial resolution of 250 m. The advantage of 250-m pixels is that less forestland is captured as fully forested pixels that are more likely to saturate pixel reflectance, and more forestland is captured within partially forested, spatially coarse pixels that reflect both forest and nonforest cover. This advantage provides a novel explanation of why modeling at coarser spatial scales improves per-pixel estimates of forest stand or canopy attributes. Studies report that errors for per-pixel estimates of forest volume and biomass decline from over 50% to 10–12% as 20 to 30-m pixels are aggregated to larger pixels of 19 ha (Reese et al., 2002) and up to 360 ha (Kennedy et al., 2000). Models of leaf area index also improve when aggregating pixels from 30 m to 500–1000 m (Cohen et al., 2003). Our own preliminary analyses revealed that biomass model correlations decreased if we increased the minimum fraction of forest area in the pixels that were included in a model.

In fact, we propose that tree or forest cover can relate to forest biomass density of a pixel in two ways. First, mass balance tells us that for uniform forest, the forest biomass density of a pixel is directly proportional to forest cover. By assuming that each pixel within the forest mask is fully forested, biomass density becomes a function of tree or forest cover for a uniform forest. Secondly, and in addition to simple mass balance, more fragmented forest or forest adjacent to nonforest (and associated with less surrounding tree or forest cover) is more likely to be disturbed or young (Helmer, 2000), have less biomass per ha of forest (Brown et al., 1993; Laurance et al., 1997), and have lower mean canopy heights (E. Helmer, unpublished data). Under this scenario, tree or forest cover data are among the most important predictor variables where forest cover is less than about 60%. A clear strength of the MODIS tree cover product, then, is that it is a global product that explains significant variance in forest biomass when data range from low to high tree cover. The weakness of proportional tree or forest cover is that these variables reach their maximums before forest biomass does. For example, the MODIS-derived tree cover product explains 37% of the variance in mean forest canopy heights across the Amazon basin where tree cover is at least 20% (N=3828), but only 1.6% of the variance in mean forest canopy heights where it is at least 60% (N=2734). Mean canopy heights for forest with at least 60% or 75% tree cover do not significantly differ (Helmer & Lefsky, 2006; E. Helmer, unpublished data).

A second potential source of per-pixel error is the spatial mismatch between the size of an FIA plot, which is distributed

over 0.67–2.5 ha (depending on region of the country), and the larger, 250-m pixels that extend over 6.25 ha. This situation, where a single FIA plot may not represent the average of the surrounding 6 ha, could inflate error estimates where local variability in forest biomass is high (for biomass estimates) and/or land cover heterogeneity is high (for forest/nonforest estimates). If so, the model errors from these site-specific assessments may conservatively gauge pixel-level errors in biomass densities. This effect of spatial mismatch on model performance measures has been noted by others (Congalton & Plourde, 2000; Foody, 2002; Smith et al., 2003; Verbyla & Hammond, 1995).

A third potential source of error is that pixels with less than 0.5 predicted probability of forest were considered 'nonforest' and received no biomass estimates, even though they could contain forest cover and biomass. Likewise, pixels having more than 0.5 predicted probability forest were considered forest. This tendency to underestimate forest area in sparsely forested regions, and overestimate it in heavily forested ones, is well documented for thematic land cover classifications of coarse spatial resolution pixels (Kuusela & Päivinen, 1995; Mayaux & Lambin, 1995; Nelson, 1989). Furthermore, FIA plot-based estimates pertain to forest land use, while satellite image-based estimates portray forest land cover. FIA definitions of forest land use and land cover are equivalent in many, but not all areas. For example, a change from forest cover to nonforest cover occurs when harvest, wildfire, windstorm, or other events result in removal of standing live trees. Such treeless areas still are defined as forest land use, assuming that regeneration is expected to occur and other land uses are not intended. Conversely, some areas having extensive tree cover are defined as nonforest use, due to other prevailing uses of the land, e.g., treed picnic areas, parks, and golf courses. In addition, effective differences in definition exist between what is observed and inventoried on the ground (e.g. total aboveground tree biomass; tree-covered residential areas) and what is captured by a satellite-borne optical sensor (e.g. tree biomass visible from above). Thus, some apparent discrepancies between plots and pixels, and resulting decreases in model accuracies, may, in fact, be artifacts of definitional inconsistencies between land use and land cover, and differences between ground inventory and optical satellite perspectives. Independent efforts are being initiated to assess these discrepancies, including use of non-FIA datasets for pixel accuracy and error and demographic data for differentiating land use from land cover.

Not surprisingly, a closer correspondence was observed between spatial aggregations of statewide map-based estimates and FIA plot-based estimates than between per-pixel comparisons with individual plots. These results are like those of Muukkonen and Heiskanen (2005) who reported large estimation errors of forest stand biomass, but spatially aggregated map-based estimates of forest biomass were comparable to municipality-level estimates from Finland's National Forest Inventory. With regard to scales of aggregation, model-based estimates of forest proportion and forest biomass tended to agree with plot-based estimates at all four scales tested. This is interesting in that despite sometime extremely high per-pixel percent errors for biomass, spatial aggregations can still provide reasonable estimates. This may be due to the fact that the per-pixel accuracy assessment is negatively impacted by the fact that the plot, which is taken to characterize the entire pixel, is very small in size relative to the size of the pixel, and furthermore it is only a single sample from the pixel. This negative effect is ameliorated to some degree by the "averaging effects" of the larger area of the hexagons; i.e., some of the errors from cancel each other.

For some geographic locales, however, particularly for the biomass map, hexagon aggregations with large absolute τ-values raise concerns over the utility of the map in those specific areas. This lack of consistency is not surprising, given the variability in ecological conditions, image data, and plot data across the US. Many of those states with the largest differences between the plot-based and map-based estimates of forest biomass and forest area are states where the most recent available data were from an older periodic inventory, or where data were not available statewide, or where poor GPS coordinates or other conditions made modeling particularly difficult. In addition, some of the differences observed may also reflect differences in definition (total aboveground tree biomass versus tree biomass visible from satellite-borne optical sensors). A relationship also exists between the difference in the estimates for forest land and the difference in estimates for forest biomass, implying that improvements in the initial forest/nonforest mask, or use of a different cutoff in the forest probability map, might increase compatibility between plot-and map-based estimates in some areas.

Presenting uncertainty maps in conjunction with the nationwide forest biomass map emphasizes that the biomass estimates are somewhat imprecise and that their uncertainty varies by location. It is important to include this uncertainty information in assessing the reliability of model-based estimates of forest area and biomass. The map-based estimates of nationwide total live above ground biomass yield estimates of total forest C storage that are within the range of previous map- and plot-based estimates of C storage or biomass, and they are consistent with the consensus that forests in the Northern hemisphere are a net C sink (Pacala et al., 2001; Schimel et al., 2001).

Zone discrepancies still exist in the current final map of aboveground forest biomass presented here. Considerable effort went in to compiling and screening the FIA data, however some areas were still handicapped by holes in the available data (e.g. zones 26, 32, 34, 35, and 36 in TX and OK), out-of-date plot data in an area of rapid change (much of the Southeast), and low quality GPS coordinates for the FIA plots (several states in the Southeast). These show up in the current map as distinct lines between zones where one side of the line may have been modeled with local but inaccurate data, and the other side of the line was modeled with more accurate but more distant data requiring an extrapolation of the model into the area of interest. This project, among others, highlights the important effects on mapping of both quantity and quality of FIA plot data, and the high value of improving such data. The current efforts within FIA including the shift to annual inventory, complete coverage with GPS and consistent data collection protocols nationwide

should substantially alleviate these problems for future modeling efforts.

5. Conclusions

Spatially explicit forest biomass information at the scale of the US provides an unprecedented picture of how forest biomass is distributed spatially across US landscapes and permits visual assessment of forest biomass distribution. It synthesizes point data from tens of thousands of ground plots into one spatial dataset that can easily feed into those ecosystem and atmospheric models that do not assimilate the point-based data. The accuracy assessments reflect the understanding that the data are primarily useful for coarse-scale modeling. The accompanying spatially explicit datasets of model uncertainty provide information critical to estimating uncertainty in such atmospheric, ecosystem, or other models and estimates (Brown et al., 1993; Brown & Schroeder, 1999; Canadell et al., 2000; Dong et al., 2003; Nemani et al., 2003; Potter, 1999). Nationwide spatially explicit modeling of forest characteristics with ground-based inventory data presented logistical and institutional challenges. Although overcoming those challenges required extensive national coordination, it forged an institutional process for nationwide forest attribute mapping that benefited from regional expertise.

Appendix A. Per-pixel measures of performance for forest/nonforest maps based on independent test sets, by zone within regions

Mapping zone	PCC	Kappa	Sensitivity	Specificity	Test set sample size
Northeast					
52	0.91	0.18	0.18	0.97	160
60	0.86	0.69	0.73	0.94	154
61	0.88	0.73	0.91	0.82	171
62	0.86	0.71	0.88	0.82	104
63	0.84	0.62	0.9	0.71	70
64	0.76	0.36	0.83	0.54	55
65	0.88	0.73	0.95	0.75	139
66	0.95	0.52	0.99	0.45	328
Northcentral					
30, 31	0.98	0.56	0.50	0.99	1613
33, 38, 43	0.94	0.50	0.54	0.97	848
41	0.89	0.78	0.92	0.87	754
44	0.86	0.80	0.86	0.94	696
47	0.80	0.53	0.65	0.87	221
49	0.95	0.82	0.80	0.98	299
50	0.90	0.80	0.90	0.90	632
51	0.89	0.78	0.84	0.93	735
Interior West					
10	0.92	0.59	0.99	0.49	525
12	0.94	0.72	0.75	0.97	883
13	0.96	0.50	0.58	1.00	121
14	0.93	0.10	0.08	0.99	220
15	0.83	0.63	0.88	0.75	338
16	0.75	0.42	0.92	0.47	298
17	0.89	0.69	0.74	0.94	397
Interior West					
18	0.94	0.38	0.35	0.98	413
19	0.92	0.84	0.93	0.91	410
20	0.96	0.41	0.32	0.99	535
21	0.88	0.75	0.91	0.84	245
22	0.97	0.42	0.28	1.00	466
23	0.8	0.48	0.62	0.87	332
24	0.89	0.68	0.73	0.94	486
25	0.89	0.54	0.52	0.96	366
27	0.92	0.50	0.47	0.97	287
28	0.82	0.64	0.87	0.77	292
29	0.95	0.66	0.62	0.98	582
Pacific Northwest					
1	0.89	0.63	0.95	0.63	3245
2	0.89	0.75	0.93	0.81	1832
3	0.88	0.74	0.91	0.83	1949
4	0.97	0.69	0.65	0.99	3442
5	0.99	0.00	0.00	1.00	2150
6	0.87	0.74	0.87	0.87	2406
7	0.85	0.71	0.89	0.82	2981
8	0.97	0.54	0.45	0.99	1457
9	0.87	0.54	0.56	0.94	5464
Southern					
26, 32, 36	0.96	0.86	0.86	0.99	188
34	–	–	–	–	0
35	–	–	–	–	0
37	0.88	0.73	0.93	0.79	410
45	0.88	0.66	0.65	0.96	216
46	0.81	0.43	0.90	0.59	594
47	0.87	0.71	0.78	0.83	149
48	0.85	0.69	0.93	0.74	119
53	0.87	0.64	0.93	0.69	106
54, 59	0.86	0.35	0.95	0.35	156
55	0.88	0.59	0.92	0.70	197
56	0.67	0.39	0.61	0.73	63
57	0.83	0.38	0.96	0.35	93
58	0.86	0.35	0.97	0.30	118
Alaska	0.94	0.88	0.93	0.95	6553
Puerto Rico	0.79	0.57	0.93	0.64	28

Appendix B. Per-pixel measures of performance for biomass maps based on independent test sets, by zone within region

Mapping zone	Average absolute error	Relative error	Correlation coefficient	Test set sample size
Northeast				
52	88.3	1.39	0.17	19
60	63.2	0.82	0.53	65
61	65.0	0.97	0.32	133
62	64.5	0.95	0.39	171
63	67.8	0.93	0.44	159
64	57.0	0.91	0.40	195
65	60.4	0.95	0.22	142
66	49.7	0.73	0.33	272
Northcentral				
30, 31, 29, 40, 42	43.0	0.85	0.39	51
33, 38, 43	40.8	0.98	0.35	50
44, 49	44.4	1.03	0.12	347

Appendix B (*continued*)

Mapping zone	Average absolute error	Relative error	Correlation coefficient	Test set sample size
Northcentral				
41	40.6	0.90	0.44	313
50	44.1	0.87	0.47	308
51	46.0	0.90	0.45	197
Interior West				
10	66.0	0.81	0.57	468
12, 13, 14	17.1	0.90	0.31	123
15	26.1	0.58	0.71	230
16	39.9	0.76	0.57	203
17, 18	22.4	0.86	0.48	103
19	52.0	0.84	0.46	207
20, 29	34.5	0.72	0.59	80
21	50.4	0.75	0.59	134
22, 23	24.4	0.76	0.68	105
24	17.9	0.74	0.69	114
25, 27	18.0	0.64	0.71	85
28	39.9	0.68	0.63	171
Pacific Northwest				
1	132.6	0.81	0.51	529
2	157.1	0.88	0.46	169
3	144.8	0.84	0.52	146
4, 5	93.1	0.93	0.28	45
6	119.2	0.76	0.49	267
7	116.6	0.69	0.57	535
8, 9	51.58	0.85	0.45	289
Southern				
26, 32, 36	26.8	1.02	0.20	17
34	-	-	-	0
35	-	-	-	0
37	54.7	0.96	0.26	256
45	40.7	0.86	0.35	48
46	63.5	0.93	0.25	414
47	65.5	1.05	0.14	94
48	56.6	0.96	0.24	67
53	62.5	0.89	0.31	82
54	51.7	1.06	0.08	119
55	64.6	0.95	0.28	159
56	94.6	0.74	0.39	26
57	59.7	0.97	0.38	73
58	117.6	0.89	0.34	138
59	87.2	0.96	0.39	96
Alaska	91.5	0.59	0.69	430
Puerto Rico	65.0	0.51	0.92	*

Average absolute error is reported in metric tons per hectare. *10-fold cross-validaton.

References

Bauer, E., & Kohavi, R. (1998). An empirical comparison of voting classification algorithms: bagging, boosting, and variants. *Machine Learning*, *5*, 1−38.

Bechtold, W. A., & Patterson, P. L., (Eds). (2005). *The Enhanced Forest Inventory and Analysis Program—National Sampling Design and Estimation Procedures*. General Technical Report GTR-SRS-080. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station.

Birdsey, R., & Heath, L. (1995). Carbon changes in U.S. forests. In L. Joyce (Ed.), *Productivity of America's forests and climate change* (pp. 56-70). Fort Collins, CO: USDA Forest Service General Technical Report RM-271, Rocky Mountain Forest and Range Experiment Station.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, *26*, 123−140.

Breiman, L. (1998). Arcing classifiers (with discussion). *Annals of Statistics*, *26*, 801−849.

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees.* Monterey, CA: Wadsworth and Brooks/Cole.

Brown, S. L., Iverson, L. R., & Lugo, A. (1993). Land use and biomass changes of forests in peninsular Malaysia during 1972–1982: use of GIS analysis. In V. H. Dale (Ed.), *Effects of Land Use Change on Atmospheric CO2 Concentrations: South and Southeast Asia as a Case Study* (pp. 117−143). New York: Springer-Verlag.

Brown, S., & Lugo, A. E. (1992). The storage and production of organic matter in tropical forests and their role in the global carbon cycle. *Biotropica*, *Vol. 14*, 161−187.

Brown, S. L., & Schroeder, P. E. (1999). Spatial patterns of aboveground production and mortality of woody biomass for eastern U.S. forests. *Ecological Applications*, *9*, 968−980.

Cairns, M. A., Brown, S., Helmer, E., & Baumgardner, G. (1997). Root biomass allocation in the world's upland forests. *Oecologia*, *111*, 1−11.

Canadell, J. G., Mooney, H. A., Baldocchi, D. D., Berry, J. A., Ehleringer, J. R., Field, C. B., et al. (2000). Carbon metabolism of the terrestrial biosphere: a multi-technique approach for improved understanding. *Ecosystems*, *3*, 115−130.

Chan, J. C. W., Huang, C., & DeFries, R. S. (2001). Enhanced algorithm performance for land cover classification using bagging and boosting. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(3), 693−695.

Cochran, W. G. (1977). *Sampling Techniques.* New York: Wiley.

Cohen, J. (1960). A coefficient of agreement of nominal scales. *Educational and Psychological Measurement*, *20*, 37−46.

Cohen, W. B., Maiersperger, T. K., Yang, Z., Gower, S. T., Turner, D. T., Ritts, W. D., et al. (2003). Comparisons of land cover and LAI estimates derived from ETM+ and MODIS for four sites in North America: a quality assessment of 2000/2001 provisional MODIS products. *Remote Sensing of Environment*, *88*(3), 233−255.

Congalton, R. G., & Plourde, L. C. (2000). Sampling methodology, sample placement, and other important factors in assessing the accuracy of remotely sensed forest maps. In G. B. M. Heuvelink & M. J. P. M. Lemmens (Eds.), *Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences* (pp. 117−124). The Netherlands: Delft University Press.

Daly, C. (2002). *PRISM monthly and annual precipitation and temperature for Alaska.* Climate Source LLC.

Daly, C., Helmer, E. H., & Quiñones, M. (2003). Mapping the climate of Puerto Rico, Vieques and Culebra. *International Journal of Climatology*, *23*, 1359−1381.

Daly, C., Kittel, T., McNab, A., Gibson, W., Royle, J., Nychka, D., et al. (2000). Development of a 103-year high-resolution climate data set for the conterminous United States. *Proceedings of the 12th AMS Conference on Applied Climatology* (pp. 249−252). Asheville, NC: American Meteorological Society.

Dong, J., Kaufmann, R. K., Myneni, R. B., Tucker, C. J., Kauppi, P. E., Liski, J., et al. (2003). Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources, and sinks. *Remote Sensing of Environment*, *84*, 393−410.

Edwards, T. C., Jr., Cutler, D. R., Zimmermann, N. E., Geiser, L., & Moisen, G. G. (2006). Effects of sample survey design on the accuracy of classification models in ecology. *Ecological Modelling*, *199*, 132−141.

Fan, S., Gloor, M., Mahlman, J., Pacala, S., Sarmiento, J., Takahashi, T., et al. (1998). A large terrestrial carbon sink in North America implied by atmospheric and oceanic carbon dioxide data and models. *Science*, *282*, 442−446.

Fiorella, M., & Ripple, W. J. (1993). Determining successional stage of temperate coniferous forests with Landsat satellite data. *Photogrammetric Engineering and Remote Sensing*, *59*, 239−246.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*, 185−201.

Fortin, M., Edwards, G. & Thomson, K. P. B. (1999). The role of error propagation for integrating multisource data within spatial models: the case of the DRASTIC groundwater vulnerability model. In K. Lowell A. Jaton (Eds.), Spatial Accuracy Assessment-Land Information Uncertainty in

Natural Resources (pp. 437-443). Chelsea, MI: Sleeping Bear Press/Ann Arbor Press.

Freund, Y., & Schapire, R. (1996). Experiments with a new boosting algorithm. In L. Saitta (Ed.), *Machine Learning: Proceedings of the Thirteenth International Conference* (pp. 148−156). San Fransisco, CA: Morgan Kauffman.

Gesch, D., Oimoen, M., Greenlee, S., Nelson, C., Steuck, M., & Tyler, D. (2002). The national elevation dataset. *Photogrammetric Engineering and Remote Sensing*, *68*, 5−11.

Hansen, M., DeFries, R., Townshend, J. R., Carroll, M., Dimiceli, C., & Sohlberg, R. (2003). *500 m MODIS Vegetation Continuous Fields: Tree Cover.* College Park, MD: GLCF, University of Maryland.

Helmer, E. H. (2000). The landscape ecology of secondary forest in montane Costa Rica. *Ecosystems*, *3*, 98−114.

Helmer, E. H., Cohen, W. B., & Brown, S. (2000). Mapping montane tropical forest successional stage and land use with multi-date Landsat imagery. *International Journal of Remote Sensing*, *21*, 2163−2183.

Helmer, E. H., Lefsky, M. (2006). Forest canopy heights in Amazon River basin forests as estimated with the Geoscience Laser Altimeter System (GLAS). In Aguirre-Bravo, C. Pellicane, Patrick J. Burns, Denver P. & Draggan, Sidney, eds. 2006. Monitoring Science and Technology Symposium: Unifying Knowledge for Sustainability in the Western Hemisphere. 2004 September 20-24; Denver, CO. Proceedings RMRS-P-37CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. CD-ROM.

Helmer, E. H., Ramos, O., Lopez, T. D. M., Quiñones, M., & Diaz, W. (2002). Mapping forest types and land cover of Puerto Rico, a component of the Caribbean biodiversity hotspot. *Caribbean Journal of Science*, *38*, 165−183.

Holland, E. A., Brown, S., Potter, C. S., Klooster, S. A., Fan, S., Gloor, M., et al. (1999). North American carbon sink. *Science*, *283*, 1815.

Homer, C. G., & Gallant, A. (2001). *Partitioning the conterminous United States into mapping zones for Landsat TM land cover mapping.* USGS Draft White Paper available at http://landcover.usgs.gov

Houghton, R. A. (1999). The annual net flux of carbon to the atmosphere from changes in land use 1850-1990. *Tellus*, *51B*, 298−313.

Houghton, R. A. (2003). Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000. *Tellus*, *55B*, 378−390.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., & Ferreira, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, *83*, 195−213.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., et al. (2002). An overview of MODIS Land data processing and product status. *Remote Sensing of Environment*, *83*, 3−15.

Kennedy, P., Folving, S., Estreguil, M., Rosengren, E., Tomppo, E., Pereira, J. M., et al. (2000). Forest information from remote sensing — biomass and wood volume assessment and mapping. In T. Zawila-Niedzwiecki (Ed.), *Proceedings of the 4.02.05 Group Session; Remote Sensing and World Forest Monitoring (held within IUFRO XXI World Congress on Forests and Society, pp. 17–32).* Warsaw: Institute of Geodesy and Cartography.

Kuusela, K., & Päivinen, R. (1995). On the classification of ecosystems in boreal and temperate forests. In P. J. Kennedy, R. Päivinen, & L. Roihuvuo (Eds.), *Proceedings of an International Workshop Designing a System of Nomenclature for European Forest Mapping*, (pp. 387-393). Joensuu, Finland: Institute for Remote Sensing Applications, Joint Research Centre, European Commission and the European Forest Institute.

Laurance, W., Laurance, S., Ferreira, L., Rankin-de Merona, J., Gascon, C., & Lovejoy, T. (1997). Biomass collapse in Amazonian forest fragments. *Science*, *278*, 1117−1118.

Mayaux, P., & Lambin, E. F. (1995). Estimation of tropical forest area from coarse spatial resolution data: a two-step correction function for proportional errors due to spatial aggregation. *Remote Sensing of Environment*, *53*, 1−15.

Miles, P. D., Brand, G. J., Alerich, C. L., Bednar, L. F., Woudenberg, S. W., Glover, J. F., et al. (2001). The forest inventory and analysis database: database description and users manual version 1.0. *General Technical Report NC-218.* St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station 130 pp.

Moisen, G. G., Freeman, Blackard, J., Frescino, T. S., Zimmermann, N. E., & Edwards, T. C. (2006). Predicting tree species presence in Utah: a comparison of stochastic gradient boosting, generalized additive models, and tree-based methods. *Ecological Modelling*, *199*, 176−187.

Moisen, G. G., & Frescino, T. S. (2002). Comparing five modeling techniques for predicting forest characteristics. *Ecological Modelling*, *157*, 209−225.

Mowrer, H. T. (1994). Monte Carlo techniques for propagating uncertainty through simulation models and raster-based G.I.S. In R. G. Congalton (Ed.), *Proceedings of the International Symposium on the Spatial Accuracy of Natural Resource Data Bases* (pp. 179−188). Bethesda, MD: American Society of Photogrammetry and Remote Sensing.

Muukkonen, P., & Heiskanen, J. (2005). Estimating biomass for boreal forests using ASTER satellite data combined with standwise forest inventory data. *Remote Sensing of Environment*, *99*, 434−447.

Nelson, R. (1989). Regression and ratio estimators to integrate AVHRR and MSS data. *Remote Sensing of Environment*, *30*, 201−216.

Nemani, R. R., Keeling, C. D., Hashimoto, H., Jolly, W. M., Piper, S. C., Tucker, C. J., et al. (2003). Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, *300*, 1560−1563.

Pacala, S. W., Hurtt, G. C., Baker, D., Peylin, P., Houghton, R. A., Birdsey, R. A., et al. (2001). Consistent land-and atmosphere-based U.S. carbon sink estimates. *Science*, *292*, 2316−2320.

Potter, C. S. (1999). Terrestrial biomass and the effects of deforestation on the global carbon cycle. *Bioscience*, *49*, 769−778.

Potter, C., Klooster, S., Genovese, V., & Myneni, R. (2003). Satellite data helps predict terrestrial carbon sinks. *EOS Transactions, American Geophysical Union*, *Vol. 84*, 502−508.

Prentice, I. C., Farquhar, G. D., Fasham, M. J. R., Goulden, M. L., Heimann, M., Jaramillo, V. J., et al. (2001). The carbon cycle and atmospheric carbon dioxide. In D. Yihui (Ed.), *Climate change 2001: The scientific basis* (pp. 183−237). Cambridge, UK: Cambridge University Press.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81−106.

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning.* San Mateo, CA: Morgan Kaufman Publishers.

Reese, H., Nilsson, M., Sandström, P., & Olsson, H. (2002). Applications using estimates of forest parameters derived from satellite and forest inventory data. *Computers and Electronics in Agriculture*, *37*, 37−55.

Särdnal, C. E., Swensson, B., & Wretman, J. H. (1992). *Model Assisted Survey Sampling.* New York: Springer-Verlag.

Schimel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peyline, P., et al. (2001). Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature*, *414*, 169−172.

Scott, C. T., Bechtold, W. A., Reams, G. A., Smith, W. D., Hansen, M. H., & Moisen, G. G. (2005). Sample-based estimators used by the forest inventory and analysis national information management system. In W. A. Bechtold P.L. Patterson (Eds.), *The Enhanced Forest Inventory and Analysis Program—National Sampling Design and Estimation Procedures. Gen. Tech. Rep. SRS-80*. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station.

Smith, J. H., Stehman, S. V., Wickham, J. D., & Yang, L. M. (2003). Effects of landscape characteristics on land-cover class accuracy. *Remote Sensing of Environment*, *84*, 342−349.

Thompson, M. E. (1997). *Theory of Sample Surveys.* London: Chapman and Hall.

Turner, D., Koerper, G., Harmon, M., & Lee, J. (1995). A carbon budget for forests of the conterminous United States. *Ecological Applications*, *5*, 421−436.

Valliant, R., Dorfman, A. H., & Royall, R. M. (2001). *Finite population sampling and inference: a prediction approach.* New York: Wiley.

Verbyla, D. L., & Hammond, T. O. (1995). Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. *International Journal of Remote Sensing*, *16*, 581−587.

Vermote, E. F., & Vermueulen, A. (1999). *Atmospheric correction algorithm: spectral reflectances (MOD09).* College Park, MD: University of Maryland.

Vogelmann, J. E., Howard, S., Yang, L., Larson, C., Wylie, B., & Van Driel, N. (2001). Completion of the 1990s National Land cover Data set for the

conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, *67*, 650−661.

White, D., Kimerling, A. J., & Overton, W. S. (1992). Cartographic and geometric components of a global sampling design for environmental monitoring. *Cartography and Geographic Information Systems*, *19*(1), 5−22.

Woodbury, P. B., Smith, J. E., Weinstein, D. A., & Laurence, J. A. (1998). Assessing potential climate change effects on loblolly pine growth: A probabilistic regional modeling approach. *Forest Ecology and Management*, *107*, 99−116.